UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Probability and Computational Statistics M

2627-1-F8206B003

Obiettivi formativi

Il corso è articolato in due moduli:

1. Probabilità applicata
2. Statistica computazionale

Il primo modulo di Probabilità applicata intende fornire agli studenti avanzati strumenti probabilistici necessari per comprendere la metodologia statistica e saperla applicare in ambito economico, aziendale e finanziario. Alla fine del modulo, gli studenti saranno in grado di utilizzare gli strumenti probabilistici per comprendere: le metodologie avanzate di analisi statistica multivariata, i metodi avanzati di inferenza statistica in ambito classico e bayesiano, i modelli di statistica spaziale, i modelli per dati di elevata dimensione ed i piani di campionamento complessi.
Il secondo modulo di Statistica computazionale si propone di fornire le conoscenze per lo sviluppo di tecniche computazionali per l'inferenza in modelli statistici. Verranno forniti quindi gli elementi essenziali della programmazione con R per l'implementazione di tali tecniche.
Per tutti i dettagli si vedano i syllabi dei singoli moduli del corso.

Contenuti sintetici

MODULO 1. Dopo un'introduzione alle diverse definizioni di probabilità, verranno presentate le basi della teoria assiomatica di Kolmogorov su cui si poggia la probabilità moderna. Verranno analizzate le proprietà elementari della probabilità, tra cui la continuità, sub-additività, monotonia, inoltre verranno presentati i Lemmi di Borel-Cantelli.
Ampio spazio verrà dato ai vettori aleatori negli spazi euclidei n-dimensionali ed alle trasformazioni di vettori aleatori. Il concetto di valore atteso condizionato sarà definito ed analizzato in dettaglio, con qualche cenno alla teoria della misura.
Nella seconda parte del corso, verranno studiati i quattro concetti di convergenza di variabili aleatorie: in

distribuzione, in probabilità, quasi certa e in media r-esima. Saranno quindi presentati e dimostrati i teoremi limite del calcolo delle probabilità e le loro conseguenze.
Infine saranno definiti ed analizzati i vettori Gaussiani attraverso la funzione caratteristica.
Il corso sarà affiancato da molti esercizi pratici.

MODULO 2. Il secondo modulo affronterà diversi argomenti di statistica computazionale. Definizione di numeri casuali e pseudo-casuali. Algoritmi per la generazione di numeri pseudo casuali, test di casualità. Introduzione al metodo Monte Carlo e al principio plug-in. Introduzione ai metodi di ricampionamento jackknife e bootstrap. Aspetti numerici e grafici per l'analisi di verosimiglianza.

Programma esteso

MODULO 1.

1. **INTRODUZIONE**. Cenni storici al calcolo delle probabilità: i problemi classici. Definizioni della probabilità: classica, soggettiva e frequentista. Il principio di coerenza di Bruno de Finetti e le sue conseguenze. L'assiomatizzazione della probabilità di Kolmogorov.
2. **ASSIMI DELLA PROBABILITA' E CONSEGUENZE**. La definizione assiomatica di probabilità. Le implicazioni della definizione: additività, monotonia, disuguagliaza di Boole, continuità della probabilità. I lemmi di Borel-Cantelli. Le probabilità condizionate e l'indipendenza di eventi.
3. **VARIABILI ALEATORIE E VETTORI ALEATORI**. Definizione di variabile aleatorie e vettore aleatorio (discreti e continui). Il concetto di distribuzione e la funzione di ripartizione. Relazioni tra variabili aleatorie: condizionamento ed indipendenza. Trasformazioni di vettori aleatori: il teorema del diffeomorfismo.
4. **VALORI ATTESI**. Richiami su speranza matematica, varianza e covarianza. La disuguaglianza di Markov. Valore atteso condizionato e sue proprietà.
5. **CENNI DI TEORIA DELLA MISURA**. La probabiltà come misura. Le variabili aleatorie nella teoria della misura. L'integrale alla Lebesgue ed il lavore atteso. Definizione generale di valore atteso condizionato data una sigma-algebra (cenni).
6. **CONVERGENZA DI VARIABILI ALETORIE**. La convergenza delle variabili aleatorie: in distribuzione, in probabilità, in media r-esima e quasi certa. Le relazioni tra le diverse convergenze. Legge debole dei grandi numeri, cenni alla legge forte di Kolmogorov.
7. **FUNZIONI GENERATRICI**. Funzione caratteristica e generatrice dei momenti. Il teorema di continuità di Lévy. Il teorema centrale di convergenza. Il metodo delta.
8. **VETTORI GAUSSIANI**. Funzione caratteristica per vettori. I vettori Gaussiani.

MODULO 2.

1. Algoritmi per la generazione di numeri pseudocasuali: tecniche di inversione della funzione di ripartizione, algoritmo accettazione-rifiuto, metodi basati su trasformazioni di variabili casuali, metodi composti, rapporto di uniformi.
2. Test di casualità.
3. Introduzione al metodo Monte Carlo.
4. Metodi di riduzione della varianza dello stimatore Monte Carlo: il metodo delle variabili di controllo e il metodo delle variabili antitetiche.
5. Metodi di ricampionamento: il bootstrap e il jackknife.
6. Intervalli di confidenza bootstrap.
7. Cenni alla verifica d'ipotesi in ambito bootstrap.

Prerequisiti

Per il primo modulo è richiesta la conoscenza degli argomenti trattati nei corsi di Calcolo delle probabilità e Analisi matematica (I e II) a livello di Laurea triennale, mentre non sono previste delle propedeuticità formali per il secondo, pur essendo auspicabile una conoscenza di base dell'inferenza statistica, del calcolo delle probabilità e del linguaggio R.

Metodi didattici

I metodi didattici relativi ai due moduli sono indicati nei rispettivi syllabus.

Modalità di verifica dell'apprendimento

MODULO 1. L'esame è costituito da una prova scritta, l'orale è facoltativo. La prova scritta è costituita da esercizi e da alcune domande di teoria. Gli esercizi mirano ad accertare la comprensione degli argomenti trattati e la capacità dello studente di applicare i concetti della probabilità. Le domande di teoria servono a verificare la conoscenza e la comprensione dei concetti della probabilità. Le domande di teoria possono riguardare anche dimostrazioni svolte durante il corso.
L'orale è facoltativo e può essere chiesto sia dallo studente che dal docente. L'esame orale verte su tutto il programma del corso e deve essere svolto pochi giorni dopo lo scritto, in base alle disponibilità del docente. In tal caso il voto finale è una media della prova scritta e della prova orale. Nel caso di scritto svolto a distanza, per ragioni legate al Covid, l'orale è obbligatorio.
Durante lo scritto è consentito l'uso della calcolatrice scientifica, ma non è ammesso l'uso di appunti, libri e strumenti tecnologici. In emergenza Covid le prove scritte e orali si terranno attraverso la piattaforma Webex ed esamionline

MODULO 2.
Studenti frequentanti: esame scritto e parte computazionale con R.
Studenti non frequentanti: esame scritto e parte computazionale con R.
Durante l'esame sarà valutata la correttezza e la chiarezza delle risposte. L'esame mira a valutare le competenze descritte negli obiettivi formativi.
L'esame scritto consta di 3 domande a risposta aperta che includono domande teoriche ed esercizi da svolgere con R/Rstudio attraverso la Piattaforma degli Esami Informatizzati.
Studenti e studentesse, così come il docente, possono richiedere una prova orale facoltativa riguardante l'intero programma.
Durante la prova non è ammesso l'uso di testi o altro materiale con l'esclusione dei codici che verranno messi a disposizione dal docente all'inizio della prova.
Durante la prova non è ammesso l'uso del cellulare, né di alcun supporto digitale.

Testi di riferimento

MODULO 1.
Testo consigliato (con esercizi):

- G. Dall'Aglio (2003). Calcolo delle Probabilità. Zanichelli, terza edizione.
  Testi di consultazione:
- Grimmett G. and Stirzaker D. (2001). Probability and random processes. Oxford University Press.
  Eserciziari:

- Epifali, I. e Ladelli, L. (2021). Esercizi di probabilità per l'ingegneria, le scienze e l'economia. Edizioni La Dotta.
- Grimmett G. and Stirzaker D. (2000). One Thousand Exercises in Probability: Third Edition. Oxford Univseristy Press.

MODULO 2.

- Appunti delle lezioni a cura del docente del corso.
- Letture consigliate per integrare le lezioni:
- Robert, C.P. e Casella, G. (2009), Introducing Monte Carlo Methods with R, New York: Springer-Verlag
- Davison and Hinkley (1997). Bootstrap Methods and their Applications, Chapman and Hall.

Periodo di erogazione dell'insegnamento

Il corso è erogato nel primo semestre.

Lingua di insegnamento

Italiano

Sustainable Development Goals

ISTRUZIONE DI QUALITÁ